Exhibit 4

An Expanded Sourcebook

Qualitative Data Analysis

Second Edition

Matthew B. Miles
A. Michael Huberman

SAGE Publications
International Educational and Professional Publisher
Thousand Oaks London New Delhi

ODTÜ KÜTÜPHANESİ
METU LIBRARY

H62
M437
1994

Copyright © 1994 by Matthew B. Miles and A. Michael Huberman

All rights reserved. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

For information address:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
6 Bonhill Street
London EC2A 4PU
United Kingdom

SAGE Publications India Pvt. Ltd.
M-32 Market
Greater Kailash I
New Delhi 110 048 India

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Miles, Matthew B.
Qualitative data analysis : an expanded sourcebook / Matthew B. Miles, A. Michael Huberman. — 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-8039-4653-8 (cl) — ISBN 0-8039-5540-5 (pb)
1. Social sciences—Research. 2. Education—Research.
I. Huberman, A. M. II. Title.
H62.M437 1994
300'.723—dc20 93-41204
CIP

04 05 17

Sage Production Editor: Rebecca Holland

METU LIBRARY
0050363020

Contents

List of Figures, Tables, and Boxes viii

Acknowledgments xii

1. Introduction 1
A. The General Problem 1
B. The Nature of This Book 2
Audiences 3
Approach 3
C. Our Orientation 4
D. Varieties of Qualitative Research 5
Recurring Features of Qualitative Research 5
Three Approaches to Qualitative Data Analysis 8
Analytic Methods: Some Common Features 9
E. The Nature of Qualitative Data 9
General Nature 9
Some Underlying Issues 9
Strengths of Qualitative Data 10
F. Our View of Qualitative Analysis 10
Data Reduction 10
Data Display 11
Conclusion Drawing and Verification 11
G. Using This Book 12
Overview 12
Format of Specific Methods 12
Suggestions for Users 12

2. Focusing and Bounding the Collection of Data: The Substantive Start 16
Tight Versus Loose: Some Trade-offs 16
A. Building a Conceptual Framework 18
B. Formulating Research Questions 22
C. Defining the Case: Bounding the Territory 25
D. Sampling: Bounding the Collection of Data 27
E. Instrumentation 34
Summary Comments 38

3. Focusing and Bounding the Collection of Data: Further Design Issues 40
A. Linking Qualitative and Quantitative Data 40
Rationale 40
Brief Description 42
Illustrations 42
Advice 43
B. Management Issues Bearing on Analysis 43
Computer Use 43
Data Management 45
Staffing and Time Planning 46
Agreements With Study Participants 47
Summary Comments 48

4. Early Steps in Analysis **50**
Our Assumptions About "Data" 51
A. Contact Summary Sheet 51
Document Summary Form 54
B. Codes and Coding 55
Reflective Remarks 66
Marginal Remarks 66
C. Pattern Coding 69
D. Memoing 72
Developing Propositions 75
E. Case Analysis Meeting 76
F. Interim Case Summary 77
Data Accounting Sheet 80
G. Vignettes 81
H. Prestructured Case 83
I. Sequential Analyses 85
Summary Comments 88

5. Within-Case Displays: Exploring and Describing **90**
Describing and Explaining 90
A. How Data Displays Work 91
B. Partially Ordered Displays 102
Context Chart 102
Checklist Matrix 105
The Transcript as Poem 110
C. Time-Ordered Displays 110
Event Listing 110
Critical Incident Chart 115
Event-State Network 115
Activity Record 117
Decision Modeling 117
Growth Gradient 118
Time-Ordered Matrix 119
D. Role-Ordered Displays 122
Role-Ordered Matrix 122
Role-by-Time Matrix 126
E. Conceptually Ordered Displays 127
Conceptually Clustered Matrix 127
Thematic Conceptual Matrix 131
Folk Taxonomy 133
Cognitive Maps 134
Effects Matrix 137
Summary Comment 141

6. Within-Case Displays: Explaining and Predicting **143**
A. Explanations and Causality 144
What Does It Mean to Explain Something? 144
A Way of Thinking About Causality 145
The Power of Qualitative Causal Analysis 147
B. Explanatory Effects Matrix 148
C. Case Dynamics Matrix 148
D. Causal Network 151
Verifying Causal Networks 163
E. Making and Testing Predictions 165
Summary Comments 170

7. Cross-Case Displays: Exploring and Describing **172**
A. Cross-Case Analysis: Overview 172
Why Do Cross-Case Analysis? 173
A Key Distinction: Variables Versus Cases 173
Strategies for Cross-Case Analysis 174
B. Partially Ordered Displays 177
Partially-Ordered Meta-Matrtix 177
C. Conceptually Ordered Displays 183
Content-Analytic Summary Table 183
Substructing a Variable 184
Construct Table 184
Decision Tree Modeling 185
D. Case-Ordered Displays 187
Case-Ordered Descriptive Meta-Matrix 187
Ordering Cases Through Summed Indices 193
Two-Variable Case-Ordered Matrix 195
Contrast Table 195
Scatterplot 197
E. Time-Ordered Displays 200
Time-Ordered Meta-Matrix 200
Scatterplots Over Time 203
Composite Sequence Analysis 204
Summary Comments 205

8. Cross-Case Displays: Ordering and Explaining **207**
A. Explaining Through Comparative Analysis 207
The Main Issue 207
Working Principles 207
B. Case-Ordered Effects Matrix 208
C. Case-Ordered Predictor-Outcome Matrix 213
Predictor-Outcome Consequences Matrix 217
D. Variable-by-Variable Matrix 219
E. Causal Models 222
Causal Chains 227
F. Causal Networks—Cross-Case Analysis 228
Antecedents Matrix 233
Summary Comments 237

9. Matrix Displays: Some Rules of Thumb **239**
A. Building Matrix Displays 240
Matrix Elements 240
Rules of Thumb for Matrix Building 241
B. Entering Matrix Data 241
Rules of Thumb for Data Entry 241
C. Drawing Conclusions From Matrix Data 242
Rules of Thumb for Conclusion Drawing 242

10. Making Good Sense: Drawing and Verifying Conclusions **245**
A. Tactics for Generating Meaning 245
1. Noting Patterns, Themes 246
2. Seeing Plausibility 246

3. Clustering 248
4. Making Metaphors 250
5. Counting 252
6. Making Contrasts/Comparisons 254
7. Partitioning Variables 254
8. Subsuming Particulars Into the General 255
9. Factoring 256
10. Noting Relations Between Variables 257
11. Finding Intervening Variables 258
12. Building a Logical Chain of Evidence 260
13. Making Conceptual/Theoretical Coherence 261
B. Tactics for Testing or Confirming Findings 262
1. Checking for Representativeness 263
2. Checking for Researcher Effects 265
3. Triangulating 266
4. Weighting the Evidence 267
5. Checking the Meaning of Outliers 269
6. Using Extreme Cases 270
7. Following Up Surprises 270
8. Looking for Negative Evidence 271
9. Making If-Then Tests 271
10. Ruling Out Spurious Relations 272
11. Replicating a Finding 273
12. Checking Out Rival Explanations 274
13. Getting Feedback From Informants 275
C. Standards for the Quality of Conclusions 277
Objectivity/Confirmability 278
Reliability/Dependability/Auditability 278
Internal Validity/Credibility/Authenticity 278
External Validity/Transferability/Fittingness 279
Utilization/Application/Action Orientation 280
D. Documentation 280
Summary Comments 286

11. Ethical Issues in Analysis 288
Some Framing Remarks 288
Ethical Theories 289
Specific Ethical Issues 290
Conflicts, Dilemmas, and Trade-offs 295
Advice 296
Summary Comments 297

12. Producing Reports 298
The General Problem 298
Audiences and Effects 299
Voices, Genres, and Stances 300
Style 301
Formats and Structures 301
Using Reports 304
Summary Comments 306

13. Concluding Remarks 307
Seeing Qualitative Analysis Whole 307
Reflections 309
Advice 310

Appendix: Choosing Computer Programs for Qualitative Data Analysis 311
1. Overview and Suggestions for Use 311
2. Software Types and Functions 311
General Types 311
Specific Functions: What to Look For 312
Other Considerations 313
3. How to Choose Software: Key Questions 313
What Kind of Computer User Are You? 313
What Kind of Database and Project Is It? 314
What Kind of Analysis Is Anticipated? 314
4. Program Characteristics 315
5. Program Developers and Distributors 315
6. References 317

References 318

Author Index 331

Subject Index 334

About the Authors 337

exploratory mode, it's as if we are trying to solve an unstated or ambiguous problem, which has to be framed and reframed as we go. Wolcott (1992) talks about this as the "theory first" or "theory after" approach. Both are workable.

So what is theory? We can talk about *implicit* theory, the preconceptions, biases, values, frames, and rhetorical habits that (for example) lead us or the people we are studying to refer to a situation as a "single-parent family"—or to the same situation as a "broken home" (Rein & Schon, 1977). In this case the implicit theories are each biased. The first puts a certain kind of human group into a class called "family" and implicitly asserts its right to exist. The second classifies the situation as imperfect or damaged and implies that "a home" must have more than one parent.

We can speak also of *explicit* theory: a set of concepts that might be organized in list form, or in a hierarchy, or in a network of propositional statements. As Wolcott (1992) notes, these might be at the level of "grand theory" (e.g., symbolic interactionism) or, more typically, include modest middle-range concepts such as *culture* or *commitment* or *innovation adoption.* Conceptual frameworks, as we describe them in Chapter 2, are the researcher's first cut at making some explicit theoretical statements.

One more distinction, to which we alluded before, seems important. Efforts to describe and explain may be "paradigmatic" (Maxwell, 1992a; Maxwell & Miller, 1992)—they involve a *variable-oriented* approach (Ragin, 1987) that deals essentially with the relations among well-defined concepts. For example, you might study adolescents' decisions to attend college by examining the relationships among variables such as socioeconomic class, parental expectations, school grades, and peer support.

Or studies may be "syntagmatic," or *process-oriented,* following the events in a specified *case* context over time (Maxwell, 1992a; Mohr, 1982). For example, you might study a particular adolescent, Nynke van der Molen, over a period of several months to follow the events and conditions (e.g., a poignant discussion with her mother on why she had never worked outside the home; Jane's experience in dissecting a frog) that were related to her decision to go to veterinary school. (For further discussion of this distinction, see Chapter 7, section A, on cross-case analysis.)

We concur with these authors that *both* approaches need to be combined for careful description and explanation. This chapter includes methods of both sorts.

From describing to explaining: The analytic progression. Usually it is hard to explain something satisfactorily until you understand just what the something is. Thus a natural progression, as Rein and Schon (1977) suggest, is from telling a first "story" about a specified situation (what happened, and then what happened?), to constructing a "map" (formalizing the elements of the story, locating key variables), to building a theory or model (how the variables are connected, how they influence each other). We have constructed a deeper story, in effect, that is both variable-oriented and process-oriented.

The progression is a sort of "ladder of abstraction" (Carney, 1990; see Figure 5.1). You begin with a text, trying out coding categories on it, then moving to identify themes and trends, and then to testing hunches and findings, aiming first to delineate the "deep structure" and then to integrate the data into an explanatory framework. In this sense we can speak of "data transformation" as information is condensed, clustered, sorted, and linked over time (Gherardi & Turner, 1987).

Naturally there is no clear or clean boundary between describing and explaining; the researcher typically moves through a series of analysis episodes that condense more and more data into a more and more coherent understanding of what, how, and why. In this chapter we mainly emphasize description, but point forward to more explanatory methods in Chapter 6.

A. How Data Displays Work

The idea of a display is central to this book. By *display* we mean a visual format that presents information systematically, so the user can draw valid conclusions and take needed action.

For qualitative researchers, the typical mode of display has been extended, unreduced text, usually in the form of written-up field notes, which the analyst scans through, attaching codes and then extracting coded segments and drawing conclusions. The analyst then writes a second form of extended text: a case study report.

Our experience tells us that extended, unreduced text alone is a weak and cumbersome form of display. It is hard on analysts because it is *dispersed* over many pages and is not easy to see as a whole. It is *sequential* rather than simultaneous, making it difficult to look at two or three variables at once. It is usually *poorly ordered,* and it can get very *bulky,* monotonously overloading. Comparing several extended texts carefully is very difficult.

The same objections apply with even stronger force for final readers. Indeed, it is been claimed (Mulhauser, 1975) that long case studies are almost useless for policymakers, who cannot afford the time required to comprehend a long account and to draw conclusions for their work.[1]

The argument of this book is, You know what you display. Valid analysis requires, and is driven by, displays that are focused enough to permit a viewing of a full data set

sions about their work; he concludes skeptically that written feedback is unwise, doesn't correct mistakes, doesn't help cooperation or empowerment, and may even cause self-censorship by the researcher.

General implication: It won't be easy. It helps to remember that data feedback is an occasion to learn more about the case, not only about your feedback.

Suggestions for practice. Early in your study, consider alternative approaches to verification through member feedback, ranging from none at all to early/intermittent to late and well-structured (case critique, verification or prediction forms, etc.); consider costs and benefits of each. For promising candidates, use several colleagues familiar with the *setting* you are studying, but not your research, to act as surrogates for real informants in tryouts of your feedback plans.

C. Standards for the Quality of Conclusions

We've reviewed 26 tactics for drawing and for verifying conclusions. How will you, or anyone else, know whether the finally emerging findings are *good?* That term has many possible definitions: possibly or probably true, reliable, valid, dependable, reasonable, confirmable, credible, useful, compelling, significant, empowering (add others of your choice). It's not enough to say that well-carried-out tactics will make for good conclusions. As Firestone (1990) notes, "The major justification for the research enterprise is that we have the time and the skills to develop approximations of the truth that have a firmer warrant than common sense" (p. 123).

As we were writing this book, we collected survey data from a wide range of qualitative researchers, asking about key analysis issues they were concerned about. The basic question of the goodness of conclusions can be approached from a "justificatory" point of view, as one of our respondents noted:

> How to really convince reductionists/positivists that naturalistic inquiry/qualitative research methods are *not* any more biased or inaccurate or imprecise than their methods? (From a specialist in recreation and leisure studies)

In this frame the issue is legitimacy. But as many researchers have pointed out, the problem of quality in qualitative studies deserves attention in its *own* terms, not just as a justification device. Other respondents said:

> It is probably an issue for all of us: when your own judgment has to be relied on for strength, significance of a finding. This is what makes the process seem subjective. (Public health researcher)

> Creation of lenses to understand my data. It is difficult to avoid seeing what you expect to see without using careful procedures. (Industrial engineer)

> If criteria for "goodness" are a false hope, how do we share/teach/defend context-bound, hermeneutically analyzed, interpretive inquiry? (Philosopher)

In this section we explore some practical standards that could help us all judge the quality of conclusions.

The battles in this domain have been extensive, and they continue. Many interpretivist researchers take the position that there is no "fact of the matter" (e.g., Schwandt, 1990) and suggest by extension that it is not really possible to specify criteria for good qualitative work—and that the effort to do so is somehow expert-centered and exclusionary, not responsive to the contingent, contextual, personally interpretive nature of any qualitative study.[22]

But the problem of quality, of trustworthiness, of authenticity of findings will not go away. The fact is that some accounts are better than others. Although we may acknowledge that "getting it all right" is an unworkable aim, we should, as Wolcott (1990a) suggests, try to "not get it all wrong." Would you be bothered if a journalist did not tell the truth about you, or if a court dismissed a case against someone who had assaulted you, saying that everyone has different interpretations?

Our view is that qualitative studies take place in a real social world, and can have real consequences in people's lives; that there is a reasonable view of "what happened" in any particular situation (including what was believed, interpreted, etc.); and that we who render accounts of it can do so well or poorly, and should not consider our work unjudgable. In other words, shared standards are worth striving for (Howe & Eisenhart, 1990; Williams, 1986).

We cannot enter here into a discussion of how "goodness criteria" flow from epistemological positions (for thoughtful discussions, see J. K. Smith, 1990, and Hammersley, 1992). Rather, we remain broadly in the "critical realist" tradition and discuss five main, somewhat overlapping, issues: the objectivity/confirmability of qualitative work; reliability/dependability/auditability; internal validity/credibility/authenticity; external validity/transferability/fittingness; and utilization/application/action orientation. (Here we are pairing traditional terms with those proposed as more viable alternatives for assessing the "trustworthiness" and "authenticity" of naturalistic research [Guba & Lincoln, 1981; Lincoln, 1990; Lincoln & Guba, 1985]).

In each section we describe the issues generally, without trying to straighten out all of the thorny problems involved. (Better people than us are still trying.) Then we propose some practical guidelines in the form of questions that can be applied to qualitative work—your own or that of others.

As in Marshall's (1990) practical list, these are not "rules" to be stiffly applied, but queries we think can be fruitfully posed when you are reflecting on the question, "How good is this piece of work?"—whether it is yours or someone else's. We cite useful sources as we go.

Objectivity/Confirmability

The basic issue here can be framed as one of relative neutrality and reasonable freedom from unacknowledged researcher biases—at the minimum, explicitness about the inevitable biases that exist. In short, do the conclusions depend on "the subjects and conditions of the inquiry," rather than on the inquirer (Guba & Lincoln, 1981)? This domain is sometimes labeled "external reliability," with emphasis on the replicability of a study by others (LeCompte & Goetz, 1982).

Relevant queries. Some useful questions to be asked of a qualitative study about this issue are:

1. Are the study's general methods and procedures described explicitly and in detail: Do we feel that we have a complete picture, including "backstage" information?
2. Can we follow the actual sequence of how data were collected, processed, condensed/transformed, and displayed for specific conclusion drawing?
3. Are the conclusions explicitly linked with exhibits of condensed/displayed data?
4. Is there a record of the study's methods and procedures, detailed enough to be followed as an "audit trail" (Schwandt & Halpern, 1988)?
5. Has the researcher been explicit and as self-aware as possible about personal assumptions, values and biases, affective states—and how they may have come into play during the study?
6. Were competing hypotheses or rival conclusions really considered? At what point in the study? Do other rival conclusions seem plausible?
7. Are study data retained and available for reanalysis by others?

Reliability/Dependability/Auditability

The underlying issue here is whether the process of the study is consistent, reasonably stable over time and across researchers and methods. We can, in effect, speak of "quality control" (Goetz & LeCompte, 1984; A. G. Smith & Robbins, 1984). Have things been done with reasonable care?

Kirk and Miller (1986) helpfully distinguish "diachronic" reliability (stability of observations over time) from "synchronic" reliability (stability in the same time frame) and point to the danger of "quixotic" reliability (what happens when multiple respondents give a monolithic, party-line answer: Remember that a broken thermometer is 100% reliable—but not very valid).

Relevant queries. What can be usefully asked in this domain?

1. Are the research questions clear, and are the features of the study design congruent with them?
2. Is the researcher's role and status within the site explicitly described?
3. Do findings show meaningful parallelism across data sources (informants, contexts, times)?
4. Are basic paradigms and analytic constructs clearly specified? (Reliability depends, in part, on its connectedness to theory.)
5. Were data collected across the full range of appropriate settings, times, respondents, and so on suggested by the research questions?
6. If multiple field-workers are involved, do they have comparable data collection protocols?
7. Were coding checks made, and did they show adequate agreement?
8. Were data quality checks made (e.g., for bias, deceit, informant knowledgeability?)
9. Do multiple observers' accounts converge, in instances, settings, or times when they might be expected to?
10. Were any forms of peer or colleague review in place?

Internal Validity/Credibility/Authenticity

Here we arrive at the crunch question: truth value. Do the findings of the study make sense? Are they credible to the people we study and to our readers? Do we have an authentic portrait of what we were looking at?

Validity itself is not monolithic; the classic, measurement-oriented view (Brewer & Hunter, 1989) differentiates face, content, convergent, discriminant, and predictive validity. More helpfully for our purposes, Maxwell's (1992b) thoughtful review distinguishes among the types of *understanding* that may emerge from a qualitative study: descriptive (what happened in specific situations); interpretive (what it meant to the people involved); theoretical (concepts, and their relationships, used to explain actions and meanings); and evaluative (judgments of the worth or value of actions and meanings). Warner (1991) also speaks of "natural" validity—the idea that the events and settings studied are uncontrived, unmodified by the researcher's presence and actions.

Furthermore, work on the "validity" of narratives emerging from interpretive studies (Connelly & Clandinin, 1990) suggests the importance of such aspects as "appar-

ency" and "verisimilitude" (Van Maanen, 1988), "authenticity," "plausibility," and "adequacy," among others.

We should also note Kvale's (1989b) emphasis on validity as a process of checking, questioning, and theorizing, not as a strategy for establishing rule-based correspondence between our findings and the "real world." "Validation becomes the issue of choosing among competing and falsifiable explanations."

Relevant queries. Some useful possibilities are:

1. How context-rich and meaningful ("thick") are the descriptions (Denzin, 1989b; Geertz, 1973)?
2. Does the account "ring true," make sense, seem convincing or plausible, enable a "vicarious presence" for the reader?
3. Is the account rendered a comprehensive one, respecting the configuration and temporal arrangement of elements in the local context (Campbell, 1986)?
4. Did triangulation among complementary methods and data sources produce generally converging conclusions? If not, is there a coherent explanation for this (Mathison, 1988)?
5. Are the presented data well linked to the categories of prior or emerging theory? Do the measures reflect the constructs in play?
6. Are the findings internally coherent (Eisner, 1991); are concepts systematically related (Strauss & Corbin, 1990)?
7. Were rules used for confirmation of propositions, hypotheses, and so on made explicit (Miller, n. d.)?
8. Are areas of uncertainty identified? (There should be some.)
9. Was negative evidence sought for? Found? What happened then?
10. Have rival explanations been actively considered? What happened to them?
11. Have findings been replicated in other parts of the database than the one they arose from?
12. Were the conclusions considered to be accurate by original informants? If not, is there a coherent explanation for this?
13. Were any predictions made in the study, and how accurate were they?

External Validity/Transferability/Fittingness

We need to know whether the conclusions of a study have any larger import. Are they transferable to other contexts? Do they "fit" (Lincoln & Guba, 1985)? How far can they be "generalized"?

Firestone's (1993) review suggests three levels of generalization: from sample to population (less helpful for qualitative studies), analytic (theory-connected), and case-to-case transfer (see also Kennedy, 1979).

Maxwell (1992c) also speaks of "theoretical" validity, the presence of a more abstract explanation of described actions and interpreted meanings. Such explanations could be considered as "internal" validity, but they gain added power if they are connected to theoretical networks beyond the immediate study. Generalizability, he suggests, requires such connection-making, either to unstudied parts of the original case or to other cases.

Schofield (1990) usefully distinguishes generalizing to "what is" (other actual contexts), to "what may be" (sites in the forefront of some similar process) and to "what could be" (outstanding or ideal cases).

The generalizing process is far from mechanical, as Noblit and Hare (1988) note in their discussion of "meta-ethnography": It is more like translating, refuting, or synthesizing two or more studies of similar phenomena. It is careful interpretation, not just "adding up."

Relevant queries. Here we may usefully ask:

1. Are the characteristics of the original sample of persons, settings, processes (etc.) fully described enough to permit adequate comparisons with other samples?
2. Does the report examine possible threats to generalizability? Have limiting effects of sample selection, the setting, history and constructs used been discussed (LeCompte & Preissle, 1993)?
3. Is the sampling theoretically diverse enough to encourage broader applicability?
4. Does the researcher define the scope and the boundaries of reasonable generalization from the study (McGrath & Brinberg, 1983)?
5. Do the findings include enough "thick description" for readers to assess the potential transferability, appropriateness for their own settings?
6. Does a range of readers report the findings to be consistent with their own experience?
7. Are the findings congruent with, connected to, or confirmatory of prior theory?
8. Are the processes and outcomes described in conclusions generic enough to be applicable in other settings, even ones of a different nature (Becker, 1990; Bogdan & Biklen, 1992)?
9. Is the transferable *theory* from the study made explicit (Maxwell, 1992b)?
10. Have narrative sequences (plots, histories, stories) been preserved unobscured? Has a general cross-case theory using the sequences been developed (Abbott, 1992b)?
11. Does the report suggest settings where the findings could fruitfully be tested further?
12. Have the findings been replicated in other studies to assess their robustness? If not, could replication efforts be mounted easily?

Utilization/Application/Action Orientation

Even if a study's findings are "valid" and transferable, we still need to know what the study does for its participants, both researchers and researched—and for its consumers. We simply cannot avoid the question of "pragmatic validity" (Kvale, 1989a); it's an essential addition to more traditional views of "goodness."

Evaluation and policy studies in particular are supposed to lead to more intelligent action; whether or not they do, real people's lives are being affected, and large amounts of money are being spent (or misspent). As Patton (1990) notes, "The ultimate test of the credibility of an evaluation report is the response of decision-makers and information users to that report" (p. 469). Managers and consultants often rely on qualitative studies (Gummeson, 1991). Action research (Watkins, 1991) is designed to solve particular local problems through a cycle of reconnaissance, planning, action, and re-reconnaissance. Critical theorists (for good overviews see Carspecken & Apple, 1992, and Popkewitz, 1990) stress the importance of "emancipatory" research that makes people aware of inequitable or oppressed positions and empowers their corrective action.

Lincoln (1990) speaks more broadly of good qualitative research as enhancing "(a) levels of understanding and sophistication and (b) the ability of participants and stakeholders to take action during and after an inquiry and to negotiate on behalf of themselves and their own interests in the political arena."

There are also questions of ethics—Who benefits from a qualitative study, and who may be harmed? (see also Chapter 11)—and of "evaluative validity" (Maxwell, 1992b): judgments made about the worth, legitimacy, or goodness of actions or meanings.

Relevant queries. What are some fruitful probes to be made here? Our range is deliberately wide.

1. Are the findings intellectually and physically accessible to potential users? (Note Gummeson's [1991] tongue-in-cheek dictum that "a scientific report should be boring and difficult to read.")
2. Do the findings stimulate "working hypotheses" on the part of the reader as guidance for future action (Donmoyer, 1990)?
3. What is the level of usable knowledge offered? It may range from consciousness-raising and the development of insight or self-understanding to broader considerations: a theory to guide action, or policy advice. Or it may be local and specific: corrective recommendations, specific action images.
4. Do the findings have a catalyzing effect leading to specific actions (Lincoln, 1990)?
5. Do the actions taken actually help solve the local problem?
6. Have users of the findings experienced any sense of empowerment, of increased control over their lives (Lincoln & Guba, 1990)?
7. Have users of findings learned, or developed new capacities?
8. Are value-based or ethical concerns raised explicitly in the report? If not, do some exist implicitly that the researcher is not attending to?

These questions are pointers, not rules. They will get clearer as the craft of qualitative analysis advances. The way to get them clearer is to ask—and try to answer—them more frequently than we do now.

At a number of points during this discussion of goodness criteria, we alluded to the importance of careful documentation and its value as an "audit trail" that peers or other colleagues can follow. We turn to this next in more detail.

D. Documentation

The Problem

Good qualitative research, like any other research, requires careful record keeping as a way of connecting with important audiences. The first audience is the *self:* The notebooks of the molecular biologist, the industrial sociologist, or the clinical drug tester help each keep track of what was done along the way, suggest ways of improving next steps, and give reassurance about the reproducibility of the results.

The second audience is *readers* of the research reports, who need to know what was done and how, as a way of assessing the credibility of the findings. Other researchers make such judgments carefully, even obsessively. And other readers such as local informants, policymakers, managers, and the general public often raise such questions as, "Whom did you talk to, anyway?" "How do you know?" "Why are you being so negative?" and "Where did you get that recommendation?"

The third audience is a subset of the second: *other researchers* who are interested in secondary analysis of the data, want to carry out a meta-analysis or synthesis of several studies, or want to replicate the findings to strengthen or refute them.[23]

For the latter two audiences, most journals require authors of empirical studies to report on their procedures as an integral part of the article. The formats are often so familiar that the author can almost fill in the blanks when writing sections on sampling, methods, and data analysis. For quantitative studies, it's also typical to report such basic data details as means, standard deviations, derived scores and scales, marginals, zero-order correlations, and various validity and reliability coefficients.